

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

Real-time Detection of Vagus Nerve Using Ultrasound B-Mode Imaging and Neural Networks for Focused Ultrasound Neuromodulation Guidance

### Permalink

<https://escholarship.org/uc/item/50940529>

### Author

Gong, Yan

### Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Real-time Detection of Vagus Nerve Using Ultrasound B-Mode Imaging and Neural Networks for  
Focused Ultrasound Neuromodulation Guidance

A Thesis submitted in partial satisfaction of the requirements  
for the degree Master of Science

in

Electrical Engineering (Signal and Image Processing)

by

Yan Gong

Committee in charge:

Professor Ramesh Rao, Chair

Professor Mingxiong Huang

Professor Imanuel Lerman

Professor Truong Nguyen

2019

Copyright  
Yan Gong, 2019  
All rights reserved.

The Thesis of Yan Gong is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

Chair

University of California San Diego

2019

## TABLE OF CONTENTS

Signature Page.....	iii
Table of Contents.....	iv
List of Figures.....	vi
List of Tables.....	vii
Acknowledgements.....	viii
Abstract of the Thesis.....	ix
Introduction.....	1
Chapter 1 Background.....	3
1.1 Ultrasound Imaging.....	3
1.1.1 Physics.....	3
1.1.2 Basic System Block Diagram.....	5
1.1.3 Beam Formation.....	6
1.2 Neuromodulation.....	7
1.2.1 Literature Review of Neuromodulation Using Focused Ultrasound.....	8
1.2.2 Vagus Nerve Stimulation.....	8
1.2.3 Safety.....	10
1.3 Medical Image Processing and Computer-Aided Diagnosis (CAD).....	12
1.3.1 Classical Methods for Medical Image Processing.....	12
1.3.2 Deep Learning and Artificial Intelligence in CAD Applications.....	13
Chapter 2 Method	
2.1 Software Beamformer and Scan Converter Pipeline.....	14
2.2 Convolutional Neural Network for Vagus Nerve Tracking.....	19
2.2.1 Faster R-CNN Model.....	19
2.2.2 Training Data Preparation.....	20
2.2.3 Proposed Image Guidance Pipeline.....	21
Chapter 3 Results	
3.1 B-mode Image Reconstruction using Beamforming.....	22
3.1.1 Beamforming Pipeline Results: Step-to-step.....	22
3.1.2 Beamforming with Different Number of Elements.....	24

3.2 Vagus Nerve Detection from Beamformed Data Using Faster R-CNN.....	27
Chapter 4 Summary and Future Work.....	31
References.....	34

## LIST OF FIGURES

Figure 1.1: System diagram for data processing of a modern ultrasound imaging transducer.....	5
Figure 1.2: Beamforming process.....	6
Figure 1.3: Vagus nerve circuits for anti-inflammatory response.....	9
Figure 1.4: Example pipeline of feature segmentation method in B-mode image.....	12
Figure 2.1: Raw data structure in the software.....	14
Figure 2.2: Grouped elements as a sliding window for beamforming of a linear array.....	15
Figure 2.3: Faster R-CNN model with Region Proposal Network.....	19
Figure 2.4: Architecture for AlexNet and VGGNet in MATLAB.....	20
Figure 2.5: Conventional ultrasound image processing pipeline for CAD applications.....	21
Figure 2.6: Proposal processing pipeline with detection network.....	21
Figure 3.1: Step-by-step outputs from B-mode image reconstruction with beamforming.....	23
Figure 3.2: Reconstructed image under simulation mode using different number of elements for beamforming.....	24
Figure 3.3: Reconstructed image under neck scan using different number of elements for beamforming.....	25
Figure 3.4: Averaged PSNR and frame rate with different number of elements.....	26
Figure 3.5: Vagus nerve detection results from beamformed data.....	28
Figure 3.6: B-mode scan using clinical device from Philips Healthcare®.....	29
Figure 4.1: Beam steering capability to compensate movements from carotid pulsation.....	32
Figure 4.2: Robotic arm programming to compensate large-scale movements.....	33

## LIST OF TABLES

Table 3.1: PSNR and frame rate comparison with different number of elements under simulation mode.....	25
Table 3.2: PSNR and frame rate comparison with different number of elements under neck scan.....	26
Table 3.3: Performance of vagus nerve detector with two models.....	29
Table 3.4: Performance of vagus nerve detectors (inducing external training data).....	30

## ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Imanuel Lerman for his support as my PI for this whole project. His inspiration and assistance on my work has been invaluable. Through multiple drafts and many long nights, his guidance has proved to be immeasurable.

I would like to acknowledge Professor Ramesh Rao for his support as the chair of my committee. His guidance and advice throughout my thesis have been greatly helpful.

I would also like to acknowledge Dr. Rahul Singh, CEO of Farus LLC., Professor Truong Nguyen and Chen Du of Video Processing Lab at UCSD, for their great dedication on my thesis. They have provided inestimable assistance and technical support during my work.

Figure 1.3 is provided by Dr. Imanuel Lerman from the paper "Noninvasive transcutaneous vagus nerve stimulation decreases whole blood culture-derived cytokines and chemokines: a randomized, blinded, healthy control pilot trial". I. Lerman, R. Hauger, L. Sorkin, J. Proudfoot, B. Davis, A. Huang, K. Lam, B. Simon and D. G. Baker, 2016. Dr. Lerman is the primary investigator and author of the paper, and he is the advisor of this thesis author.

## ABSTRACT OF THE THESIS

Real-time Detection of Vagus Nerve Using Ultrasound B-Mode Imaging and Neural Networks for  
Focused Ultrasound Neuromodulation Guidance

by

Yan Gong

Master of Science in Electrical Engineering (Signal and Image Processing)

University of California San Diego, 2019

Professor Ramesh Rao, Chair

Vagus nerve electrical stimulation is known to improve learning processes and augment anti-inflammatory treatment [1]. Current methods include the surgically implanted cuff electrode, or the noninvasive electric field that penetrates from the skin to the nerve. The former ensures target accuracy yet can be fraught with complications due to surgical implant, whereas the latter avoids procedural complications but may result in off target neurostimulation (i.e., other neuronal stimulation besides vagus

nerve) [2]. Off target neuronal stimulation can alter cardiac function and cerebral blood flow patterns [3] [4]. Converging lines of evidence now support non-invasive high temporal and spatial resolution focused ultrasound can result in brain and peripheral nerve stimulation [5] [6] [7]. Therefore, focused ultrasound is proposed to replace electrical nerve stimulation to effectively modulate the vagus nerve as a safe and noninvasive modality. To guarantee safety and efficacious vagus nerve focused ultrasound stimulation, real-time image guidance is required to provide efficient and accurate tracking of the vagus nerve. In this work, we propose to: 1) develop real time B-mode ultrasound imaging to visualize the vagus nerve and 2) employ neural networks capable of tracking the nerve location within the view. We implement the real-time beamforming process from raw data to the reconstructed B-mode image and send it to the vagus nerve detector that is trained using convolutional neural networks. We successfully extract the location of the nerve from the beamformed data, validating the idea of real-time guidance. Future work will focus on improving the efficiency of nerve detection algorithms that are then used as feedback for an eventual autonomous closed loop focused ultrasound nerve stimulation system.

## Introduction

### - Motivation

Vagal nerve stimulation (VNS) has proved to have significant effects in neural system therapies. Recent studies have shown evidence that VNS can improve learning, such as augmenting the extinction of conditioned fear thus suppressing the symptoms of post-traumatic stress disorder [8]. Moreover, decreases of inflammatory cytokines have been observed as a response of vagus nerve stimulation, suggesting this potential anti-inflammation therapy may soon be approved to treat diseases such as Rheumatoid Arthritis [1] [9]. Invasive, (surgically implanted), vagus nerve stimulators are FDA approved and show efficacy for treatment of depression and epilepsy. Common to all implantable vagus nerve stimulator devices, is the fact that invasive vagus nerve stimulation requires surgical intervention fraught with complications related to implant as well as continuous neural inflammation (glial activation) and scarring that occurs with chronic implantation of foreign objects [10]. Therefore, there is a clear need for non-invasive technologies that stimulate the vagus nerve without the possibility of imparting neuronal damage. Current state of the art non-invasive vagus nerve stimulation techniques do not likely cause direct neuronal damage, however transcutaneous nVNS electric field can penetrate deep and may result in stimulation of off target neuronal structures such as the middle cervical ganglion (i.e. the sympathetic ganglion). Prolonged off target stimulation of the sympathetic ganglion is known to cardiac dysfunction and changes cerebral blood flow patterns [3] [4]. The work in this thesis aims to circumvent potential damage to the vagus nerve (common with surgically implanted devices) and off target effects (common with non-invasive vagus nerve stimulation) by development of a focused ultrasound therapeutic and imaging system that are *capable of safely stimulating only the vagus nerve*, without off target stimulation effects. The process of focused ultrasound peripheral nerve stimulation has recently been developed and is an attractive methodology due to its noninvasive nature and high spatial resolution.

### - Problem Statement

In order to achieve safe and effective stimulation with focused ultrasound, accurate targeting of the vagus nerve is critical. Image guidance, therefore, is required to have an accurate view of important cervical neck structures including the vagus nerve within the field of interest. However, there are unavoidable movements of the nerve during the stimulation. First, the vagus is encapsulated within the external carotid artery sheath which moves with the pulse, therefore resulting in a local movement of the vagus as well. Moreover, if the patient moves, and or rotates his or her neck the targeted vagus nerve may exit the field of view. Therefore, the image guidance needs to have capability to run accurately in real-time with automatic control system.

- Hypothesis

The work in this study aims to develop the real-time image guidance with nerve detection capability for focused ultrasound noninvasive vagus nerve stimulation. The target, vagus nerve, can be auto-detected from a B-mode scan using the real-time beamforming and reconstruction integrated with a convolutional neural network. By embedding the network into the B-mode reconstruction pipeline, real-time (with delay in millisecond scale) guidance and target movement tracking will be achieved.

## Chapter 1 Background

### 1.1 Ultrasound Imaging

The development of ultrasound imaging can be dated back to the 1940s during WWII. At that time, radar and sonar technologies were developed for marine forces, and soon the medical research community applied this pulse-echo measurement for biological tissue visualization. This marked as the start of using ultrasound to image tissues and organs for medical diagnosis, that have now expanded to include variable ultrasound imaging techniques (such as; B-mode, 3D scanning and Doppler imaging) for a variety of applications. Ultrasound has been shown to be a safe and time-efficient technology for tumor detection, cardiac disease diagnosis, pregnancy tests and other clinical examinations. Moreover, ultrasound is now widely used as an adjunct guide for minimally invasive procedures (such as use for blood draw, and in multiple interventional pain procedures which depend on accurate needle guidance), that together improve procedural accuracy, efficacy and reduce procedural complications. The following subsections will provide a comprehensive background on the basics of ultrasound imaging.

#### 1.1.1 Physics

Sound waves are the mechanism by which energy is transmitted by mechanical propagations along the medium. Forces are applied to the particles, causing them to vibrate back and forth, that results in vibration spread among neighboring medium (tissue) further propagating to form the transmission of the energy. Waves are categorized into transverse and longitudinal waves, based on the directions of particle movement and wave propagation. If the wave travels in the same direction as the particle oscillates, then it is a longitudinal wave; if it travels perpendicularly to the particle movement then it is a transverse wave. Sound waves in most media are longitudinal; bone is known to be the only biological tissue that can produce transverse or (long) sound waves [11].

Frequency of the wave is the description of how fast the particles vibrate. It is the number of cycles/vibrations within a second, denoted by the symbol  $f$  and the unit Hertz (Hz). It is one of the most

important properties for sound waves. Frequency can be interpreted as what we called “tune” in daily life, where higher tunes refer to higher frequencies. The audible range for the human ear is 20Hz to 20kHz. Period (T) is the inverse of frequency which represents the time to complete one cycle. Frequency, or period, is highly related with wavelength and sound velocity by the equation

$$c = f * \lambda = \frac{\lambda}{T}$$

where velocity c is the speed of wave propagation, and wavelength  $\lambda$  is the propagation distance in one period.

Ultrasound is characterized as sound waves with frequencies higher than our audible range (>20kHz). Although we cannot hear the wave, it turns out to be a very helpful resource to “view” things. As the wave propagates through the media and hits a boundary of a structure, part of the energy bounces back as echoes. This phenomenon is called reflection and is caused by the acoustic impedance mismatch between two media. The acoustic impedance of a medium is given by the product of sound speed c and density  $\rho$ :

$$Z = \rho * c$$

And the reflection coefficient at a perpendicular incidence surface is given by:

$$\alpha_r = \left( \frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2$$

Different media have different sound velocity c and thus different acoustic impedance. If a large impedance mismatch occurs at an interface, there will be large amounts of echoes reflecting back. Moreover, echoes that travel into deep tissue and then reflect from a deeper interface take more time to reach the surface for acquisition, resulting in a time difference of the received data. Therefore, if we send a wave towards some targets and record the echoes coming back along certain amount of time, we can convert them into a structural mapping of the targets.

These essential tissue properties constitute the basis of ultrasound imaging. Modern diagnostic ultrasound probes are typically multiple elements arranged as a linear or curved array. Each element contains a piezoelectric crystal that can produce mechanical waves when excited by an applied voltage [11]. The probe acts as both transmitter and receiver during ultrasonic scanning. First, it serves as a transmitter that sends out a pulse wave to the field of interest. The transmission stops very quickly to allow time for the previously generated wave to travel through the media. Post transmitter pulse the piezoelectric crystals detect the incoming echoes (which are mechanical waves), that are then converted to an electric signal. The transmitted echoes that are bounced back and received by the transducer then processes the recorded signals in software to produce an image representing the information of the scanned field.

### 1.1.2 Basic System Block Diagram

Extensive research has been carried out that focuses on either the source signal transmission or the received data analysis, depending on different applications. For example, B-mode gives a 2D structure map to show the tissue, vessels and organs; M-mode measures movements in depth and is commonly used in cardiac motion diagnosis; Doppler imaging can detect the blood flow velocity and indicates the direction using red (towards the sensor) and blue (away from the sensor) in the image. In this project we focused on cervical neck structures with high resolution B-mode imaging. To be specific, our work will investigate the principles for receiving and processing the data to obtain a displayable image for

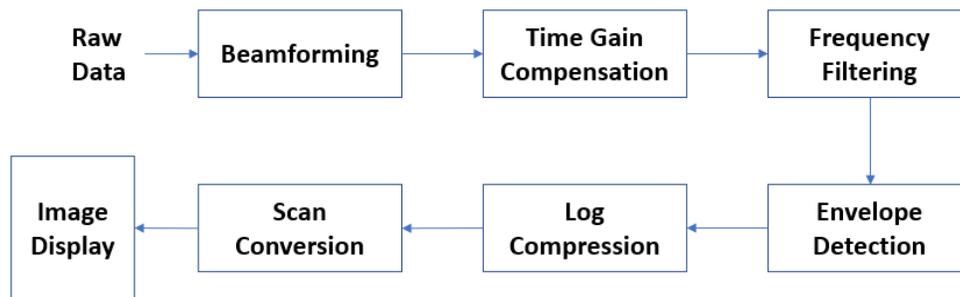


Figure 1.1: System diagram for data processing of a modern ultrasound imaging transducer

diagnosis. The processing pipeline for a modern ultrasonic imaging transducer is demonstrated as the block diagram in figure 1.1. More details on what each step does will be discussed in Chapter 2.

### 1.1.3 Beam Formation

To execute the valid signal processing for each block along the pipeline, it is important to understand the most critical step: beamforming. As the very first step in the pipeline, it ensures the correct transformation from time signals to the spatial representation as an image. The reflected echoes from the target, as time signals, travel back to the surface in radiated directions. If the elements are placed in a curved pattern pointing towards the target, then the signals will arrive at the elements at the same time. However, for a linear array this is not the case. The signals are first captured by the closest element (i.e. the one along the normal line), and then reach its neighbors and so on. Therefore, to correctly obtain the information of the target we need to compensate the time differences of signal arrival at each element. This process is called beamforming.

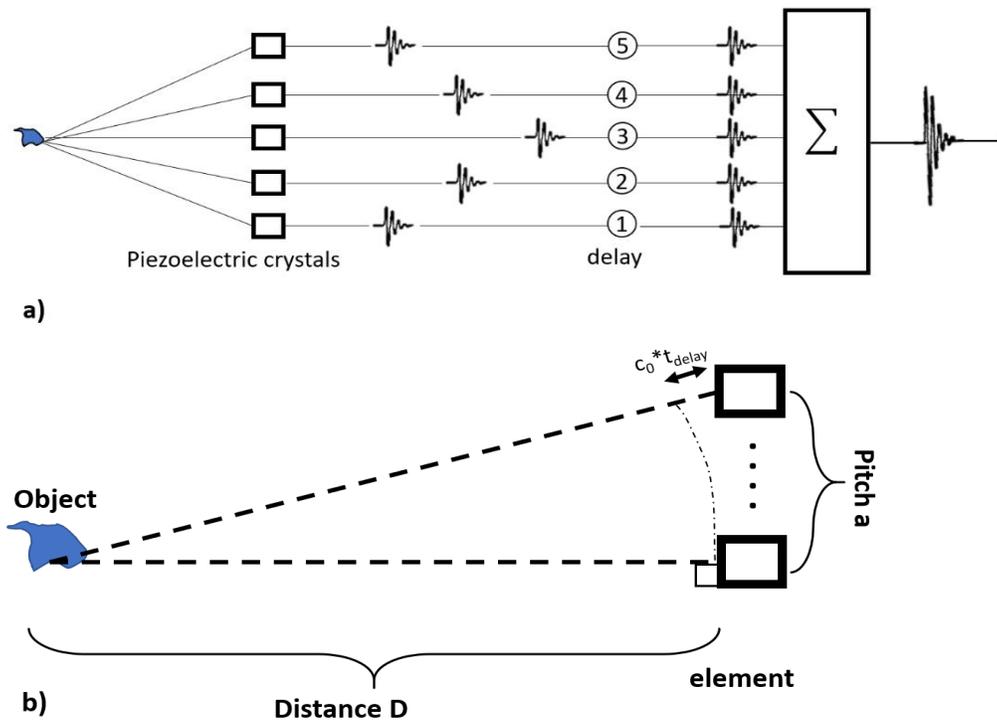


Figure 1.2: Beamforming process. a) Delay and summation configuration for a linear array, and b) Delay calculation.

Figure 1.2 demonstrates the idea of the delay process. Based on the arrival time differences at each channel, the delay process accordingly shifts the signals so that the echoes representing the same target are processed simultaneously. The relative delay time between the channels can be calculated in a simple geometry model (figure 1.2b). Suppose  $D$  (m) is the distance from the object to the probe. Consider two channels, one at the perpendicular line and the other is arbitrary. Let  $a$  (m) denotes the element pitch which is the physical distance between the center of the two elements. As we can see from the figure, the difference between the travel distances from the target can be calculated using Pythagorean Theorem. Assuming the sound velocity  $c_0$  to be constant, we can obtain the time delay between the two channels:

$$t_{delay} = \frac{1}{c_0} * (D - \sqrt{D^2 + a^2})$$

After obtaining the time delays for each element, the signals are shifted accordingly such that they become aligned without the delays. The adjusted signals are then summed up across the elements to yield one scan line, which is a more representative signal for the actual target. The summation represents the illustration in figure 1.2a, where  $r_i(t)$  is the raw data recorded by the  $i^{\text{th}}$  element and  $K$  is the number of elements to generate one scan line  $x$ . The real output of beamforming is then stacking all scan lines together, where each scan line  $x(j)$  is constructed from the local group of  $K$  elements.

$$x(j) = \sum_{i=j}^{j+K-1} r_i(t + t_{delay})$$

Beamforming is considered as the most important step for B-mode ultrasound image reconstruction, since it sets the foundation for all subsequent reconstruction steps. As the output of beamforming, the scan lines (with time delays corrected) then has the right representation for the spatial information and thus ensuring the valid image processing.

## 1.2 Neuromodulation

### 1.2.1 Literature Review of Neuromodulation Using Focused Ultrasound

Besides imaging, ultrasound has also been developed and extensively used in medical treatments for its capability to perform high resolution energy delivery, such as tissue ablation with High Intensity Focused Ultrasound (HIFU) and Focused ultrasound neuromodulation (FUS-NM). If an investigator focuses the ultrasound waves to a localized nerve area with enough intensity, the energy at that spot can modulate the neural circuit behaviors and cause crucial biological effects [5]. Recently, there is a growing interest in ultrasound neuromodulation that is known to impart significant advantages when compared to alternative neuromodulatory techniques. For example, surgically implanted deep brain stimulation electrodes are precise in terms of localization yet they require an invasive surgical operation and possible incurred procedural risks(bleeding, infection, death) [8], while transcutaneous electric field stimulation is non-invasive, but unfortunately has low spatial resolution that may cause off target effects [2] [12]. Focused ultrasound, therefore, is a competitive choice as it is; 1) a non-invasive modality, and 2) is capable of high spatial resolution that can be tuned for nerve specific neurostimulation. As early as in 1971, Gavrilov and colleagues proved the validity of high intensity focused ultrasound to stimulate peripheral nociceptors that resulted in reproducible tactile responses [6]. In 2008, Tyler and colleagues first showed FUS-NM was possible with using low intensity, low frequency focused ultrasound [13]. This seminal work showed brain stimulation FUS-NM was possible, largely due to use of low frequencies can penetrate the skull. Transcranial ultrasound stimulation modulating the somatosensory cortex has since been successfully conducted in preclinical and clinical human trials [7] [14] [15].

### 1.2.2 Vagus Nerve Stimulation

The Vagus nerve is a primary neural component of the parasympathetic nervous system, while it plays a crucial role in cardiac function, hepatic glucose production, integrated visceral regulation, as well as regulation of the inflammatory reflex [1]. With this therapeutic potential, widespread interest in modulating the vagus nerve (with both preclinical and clinical study) have been initiated to discover more efficient and effective clinical therapies. In 1997, the Food and Drug Administration approved vagus

nerve stimulation (VNS) for prevention of seizures in patients who have drug-resistant epilepsy [8]. It has been widely applied to many medical treatments such as migraine and cluster headache and is emerging as a potential treatment for Rheumatoid Arthritis [16] [17]. Carefully timed VNS can enhance neuroplasticity that is inherent to learning. For example, in a pre-clinical VNS fear study (using surgical implanted electrodes) [8], rats were exposed to a conditioned fear stimulus, and then received vagus nerve stimulation concurrent with a fear extinction paradigm. Two weeks post initial exposure, the rats were re-exposed to the conditioned fear again, while the main finding was improved fear extinction observed only in the VNS group. Similarly, patients with post-traumatic stress disorders (PTSD) whom also have dysfunctional fear extinction, may benefit from VNS. Multiple clinical trials are now underway to translate these preclinical findings. Moreover, VNS is known as a potentiate endogenous anti-inflammatory signaling, that has preliminary evidence for therapeutic treatments [1]. As shown in the model (figure 1.3a), the pathogen or tissue injury molecules activate immune cells that release inflammatory signal molecules (cytokines) that activate afferent vagal nerve fibers. An activated vagus nerve then transmits these signals to the brainstem that results in reflexive modulation of the immune cells

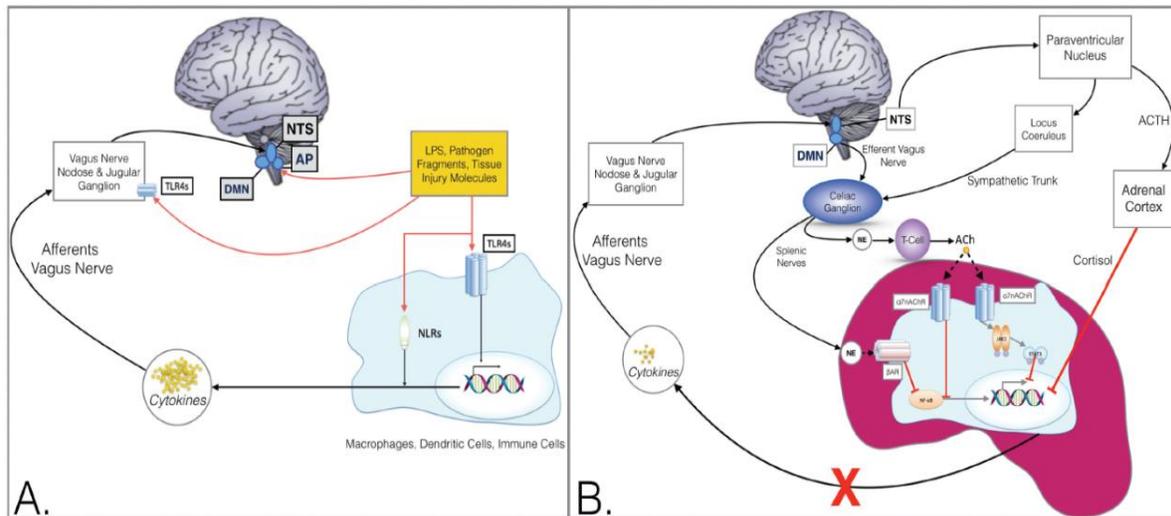


Figure 1.3: Vagus nerve circuits for anti-inflammatory response. a) Immune system gets activated and the inflammatory cytokines triggers the afferent vagus nerve, and b) Activated vagal afferents transmit from NTS to the DMN then to splenic nerve resulting in ChAT 1T cell release of the neurotransmitter acetylcholine (ACh), which tells the immune system to express the nicotinic acetylcholine receptor subunit a7 (a7nAChR). Activated a7nAChR results in suppression of the inflammatory cytokines transcription. (figure permission given by Lerman [1])

as they pass through the spleen circuit. Together the initial inflammatory insult results in eventual suppression of cytokine release and culminates in the anti-inflammatory effect (figure 1.3b). The VNS anti-inflammatory effects have been shown in multiple pre-clinical and clinical models substantiating it as a potential therapeutic target for immunomodulation.

To date all methods of vagus nerve stimulation solely employ an electrical stimulus; either with implanted cuff-electrodes or with a noninvasive electrical field. Only recently has there been an interest in vagus nerve stimulation using focused ultrasound neuromodulation. Juan and colleagues first applied focused ultrasound to the vagus nerve in 2013. They observed inhibition of the vagus nerve after continuous vagus nerve for prolonged periods of time ultrasound sonification [18]. Of particular note prolonged neuronal stimulation using focused ultrasound can result in heat and likely contributes to the reported results by Juan and colleagues. Since this initial work Irazoqui and colleagues (2019) recently show that short bursting (4ms) FUS results in modulation of the vagus nerve anti-inflammatory reflex and cause anti-inflammatory effects [19]. These promising preclinical results are clear, *but to date there is no published clinical FUS-NM to the vagus nerve*. We aim to fill this gap by developing novel, safe and non-invasive FUS-Vagus Nerve Stimulation (FUS-VNS) that may lead to a readily available clinically effective therapy.

### 1.2.3 Safety

Currently vagus nerve stimulation on human is performed through either invasive cuff electrodes positioned directly on the nerve or through the use of a noninvasive electric field. As mentioned in previous sections, implanted electrodes are invasive and require surgical operation that may result in procedural complications. Noninvasive electrical stimulation suffers from low spatial resolution and lack of specific localization of the stimulus. Therefore, focused ultrasound is a perfect choice to be both noninvasive and target specific. However, it also has its own safety concerns. Since focused ultrasound delivers a substantial amount of energy to perform neuromodulation, it is critical to ensure the safe energy transmission without causing any damage to the nerve and surrounding tissues. Too much power will

result in heat accumulation that can result in neuronal inhibition and eventual tissue damage. In addition to heat, the mechanical forces imparted by focused ultrasound to the tissue can result in microbubble formation that culminate in either stable cavitation or inertial cavitation. While stable cavitation resulting from low acoustic intensity results in small amplitude bubble vibration within very few bioeffects, inertial cavitation induces large amplitude bubble expansion that can cause tissue tearing and eventually cause organ damage [20]. Besides the control for energy input, the location of the target is also crucial as we want to avoid off target effects. Inevitable movements that occur with each pulse, as the vagus nerve is ensheathed in the carotid artery, require careful targeting and real time updating for sonification spot. Therefore, a real-time image guidance alongside the stimulation is needed to provide the target location as well as monitoring for any damage.

In many current focused ultrasound applications such as transcranial stimulation, MRI is a common technology to provide guidance for the ultrasound beam with high spatial resolution. However, MRI has high operational costs and are only available in limited locations with special experimental settings. Besides, MRI temporal resolution is inordinately slow, on the order of seconds, therefore it is incapable of real-time guidance. Ultrasound imaging, on the other hand, is mobile, wearable, provides high temporal resolution, making real-time processing possible. It has relatively high spatial resolution that is more than sufficient for medical grade diagnosis, and the algorithms to track the target (i.e. the vagus nerve). The existing diagnostic ultrasound imaging systems provide imaging power within the FDA limits, meaning that the energy from the imaging probe is safe enough not to cause any damage to the tissue.

To this date there are no systems existing that employ ultrasound imaging as guidance for focused ultrasound stimulation simultaneously, thus substantiating this study as a unique pilot. We propose to have the ultrasound imaging system to monitor the target and use the detected location to control the ultrasound stimulation system. We will integrate the real-time ultrasound imaging with a convolutional neural network for nerve tracking, which will be discussed in the next few sections.

### 1.3 Medical Image Processing and Computer-Aided Diagnosis (CAD)

#### 1.3.1 Classical Methods for Medical Image Processing

Extensive research efforts are now focused on automatic object segmentation and detection in image processing and computer vision using machine learning. For normal RGB images such as photos of traffic, researchers have built many efficient algorithms for the computer to identify pedestrians, cars, stop signs and many other different objects within the scene, making automated real-time tracking possible.

In medical image processing, there arises a similar need for such technology, namely computer-aided diagnosis (CAD). In recent years, medical researchers have been working on applying the same machine learning techniques to identify biological objects in medical images. Existing works were already reported to successfully detect certain nerves in ultrasound images or classify and differentiate abnormal tissues from healthy tissue. For example, the median nerve has been detected and auto-labeled using Median Gabor Binary Pattern as a segmentation filter combined with SVM classifier [21]. In addition, a Linear Discriminant Analysis was applied to classify tumor mass from contrast-enhanced ultrasound images [22]. The pipeline for those methods follows the classical learning flowchart (figure 1.4), including contrast enhancement, denoising, feature mapping and classification [21] [23].



Figure 1.4 Example pipeline of feature segmentation method in B-mode image

The research on CAD technology has opened huge fields of applications for machine learning in medical imaging. It can benefit doctors who aim for more efficient diagnosis with less dependency on human control and may foster future remote clinical treatment paradigms. These systems may also enable autonomous device diagnostics performed by the at home patient.

### 1.3.2 Deep Learning and Artificial Intelligence in CAD Applications

While many early-stage CAD systems relied on classical image processing pipeline and hand-craft features to recognize tissue patterns, those traditional methods failed to accommodate large variations among tissue types, image contrast, as well as between patient variability. Fortunately, the recent revolution of Artificial Intelligence has brought Deep Learning to the spotlight. Deep learning surpasses the traditional machine learning techniques with its capability to extract mid-level and high-level patterns with deeper convolutional neural networks [24]. This is extremely important for medical images, because their large variations of image resolution and large dependencies on prior medical knowledge make some tissue patterns unrecognizable by only low-level features such as edges.

There have been successful works on using deep learning methods in medical image recognition and classification. Convolutional neural networks (CNNs) were applied to X-ray images of the chest to identify pathology from the healthy ones [25]. More specific classification was achieved using CNNs on CT scans of lung to identify 7 different conditions of interstitial lung diseases [26]. Ultrasound B-mode images also gained attention in such studies due to their high temporal resolution. Of particular note, the Carotid artery (directly adjacent to the vagus nerve and a critical targeting aid in this proposal) could be detected and labeled from B-mode images using deep CNNs as classifier [27]. Similar applications worked in carotid artery plaque localization and size characterization in B-mode imaging [28]. Such capability of CNN to detect high-level tissue patterns and the fast processing time of ultrasound imaging further support CNN capability in real-time vagus nerve tracking (vagus nerve within the carotid artery sheath). We propose that we can feed the real-time beamformed data of the neck into the CNN model that computes the location of the vagus nerve, which then serves as the guidance for the control system for our FUS-VNS device.

## Chapter 2 Method

### 2.1 Software Beamformer and Scan Converter Pipeline

For the imaging transducer, we use L12-3v from Verasonics. The transducer is a 192-element linear array with center frequency of 8.92MHz. It is connected to the Verasonics Vantage System which has its compound functionalities to drive the probe and process the data. However, the received RF data goes through the built-in reconstruction functions that are opaque to customer users. In other words, we cannot directly read the line-to-line implementation of the Verasonics coding of the final B-mode image from the raw data. Verasonics does provide an “external” option, where customers can bypass their reconstruction processes with our own functions. Therefore, in the external function we implement the imaging pipeline according to the block diagram (figure 1.1). K-Wave toolbox, which is a free open source for acoustic wave simulation in MATLAB, provided a template for ultrasound linear array imaging [29]. We utilized the K-Wave template and modified it according to the specifications for the L12-3v probe and other experimental settings.

#### - Raw Data Extraction

The Raw data is received by all elements and is stored in a 2D array, called RData, with size ReceiveSample\*Elements. “ReceiveSample” represents the depth view that is tunable at the user

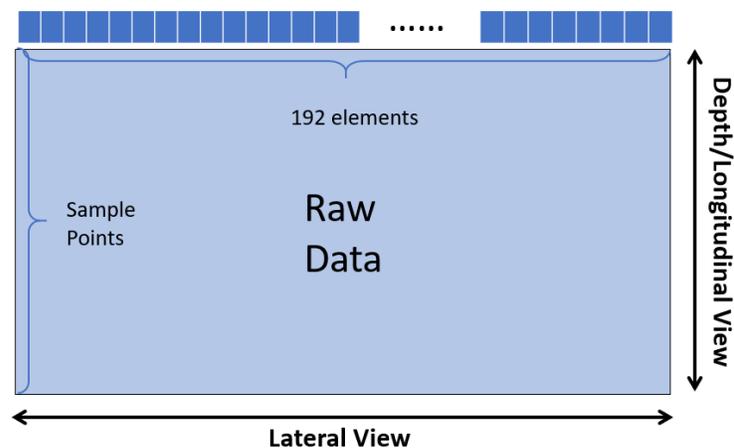


Figure 2.1 Raw data structure in the software

interface, and “Elements” are the 192 elements that cover the horizontal field of view. The RData is passed as input to the external function for reconstruction.

- Beamforming

In Raw data, each of the 192 channels doesn’t represent the vertical view (i.e. a scan line) exactly underneath that single element. Instead, one scan line consists of the information coming from a small local group of elements, as described in previous section on beam formation (see Section 1.1.3). The 1D scan line that represents the real intensity of the object is obtained by taking summation of all channels in the local group with their delays fine-tuned.

In our experiment, we compute for the complete 2D view by sliding a window of 4, 8, 16, and 32 elements through the raw data and examine the computational time cost as well as the image quality. From each group of elements, we obtain one column of scan line and move the sliding window one element to the right to calculate the next scan line.

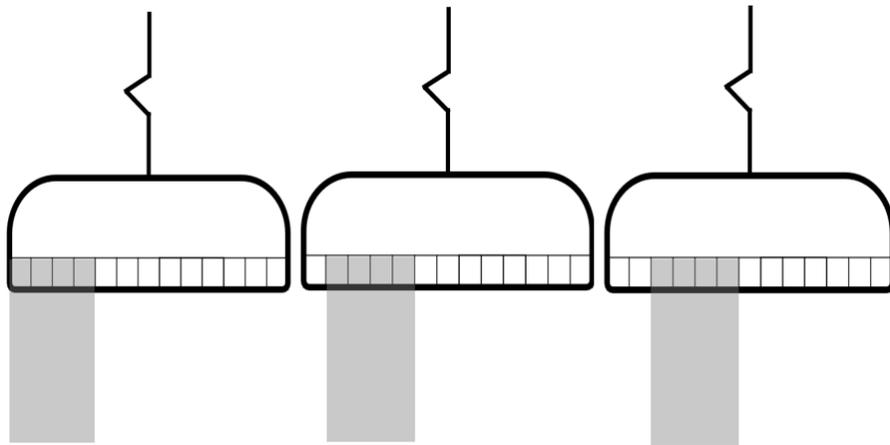


Figure 2.2 grouped elements as a sliding window for beamforming of a linear array

There is a trade-off between number of elements in a group and the spatiotemporal resolution. We can achieve higher spatial resolution if more channels contribute to one scan line. However, with more channels employed, it can take more time to compute the delay at each channel, resulting in a much

longer processing time to obtain the scan lines one by one and thus tremendously slowing down the frame rate.

- Time Gain Compensation

As the ultrasound wave travels through the tissue, the energy gets attenuated due to the interaction with biological tissue (absorption and scattering). Therefore, when a certain spot reflects an echo back to the surface, the energy signal captured by the probe is lower than the actual energy imparted at the spot. In order to compensate this energy loss, we first introduce the attenuation model which follows:

$$I = I_0 * A(d) = I_0 * e^{-\alpha f d}$$

where  $\alpha$  is the attenuation factor in dB/(MHz\*cm),  $f$  is the center frequency of the ultrasound probe,  $d$  is the total distance that the wave propagates,  $I_0$  is the original intensity, and  $I$  is the received intensity that is attenuated along  $d$ . As we can see from the formula,  $I_0$  can be recovered from  $I$  by a simple scaling:

$$I_0 = I * A(d)^{-1} = I * e^{\alpha f d}$$

- Frequency Filtering

After TGC is applied to the scan lines, a bandpass filter is applied to the data for removing the noise outside the bandwidth of the transducer. In our case, the center frequency  $f_c$  of the transducer is 8.92MHz with bandwidth 4-12Mhz. Obtaining the spectrum of the scan lines by Fourier Transform, the majority of responses fall within the transmit frequency range and some noise outside the range. In the frequency domain, we apply a Gaussian Filter centered at  $f_c=8.92$ Mhz and has 100% BW. In this way, we can preserve the signals around the fundamental frequency and smooth the noise outside our interest.

$$u_{filtered}(t) = FT^{-1}\{F_{u_{filtered}}(\omega)\} = FT^{-1}\{F_u(\omega) * G\}$$

In some applications we will want to look at harmonic imaging, this step is also useful by applying another bandpass filter to the original signal. The new filter is now centered at the first harmonic

(2fc) and has some tunable bandwidth. We eliminate this process because we don't need the harmonic image in this project, yet it remains a useful and available option if needed.

- Envelope Detection

The signals are still in modulated sinusoid pattern which has too much variation to be displayed as an image. Instead, we expect to assign the pixel value with only the peaks (maximum intensity) of the sinusoid. This is done by a demodulation process that is called envelope detection. The envelope of a signal is defined as the magnitude of its complex value:

$$E(u) = \sqrt{u_r(t)^2 + u_i(t)^2}$$

where the complex signal is defined as;  $u = u_r(t) + ju_i(t)$ . In general, the waves are complex signals with both real and imaginary parts. In our case, however, all the processes so far are based on the RF data which only stores the real signal  $u_r(t)$ . To accurately compute for the envelope  $E(u)$ , we need to reconstruct the imaginary part  $u_i(t)$  from the given real part  $u_r(t)$ . This is done by taking the Hilbert Transform  $H(\cdot)$  of the real part. In signal processing, we know that:

$$u_i(t) = H(u_r(t)) = u_r(t) ** \left(\frac{1}{\pi t}\right)$$

where the Hilbert Transform of  $u_r(t)$  is the convolution of  $u_r(t)$  and  $1/(\pi t)$ . Usually this is done in frequency domain, when the relationship between their Fourier Transform is:

$$F_{u_i(t)}(\omega) = \mathcal{F}\{H(u_r(t))\} = F_{u_r(t)} * (-j \operatorname{sgn}(\omega))$$

where  $\operatorname{sgn}(\omega)$  denotes the sign of  $\omega$ . The second term is a  $90^\circ$  phase shift based on the sign of frequency. Therefore, given the scan lines data which is the real signal, we can construct the corresponding imaginary signal using Hilbert Transform and finally obtain the envelope.

- Log Compression

The proceeded scan lines have a relatively large dynamic range that is not suitable for a gray scale display. Therefore, we need to remap the pixel intensities to a more reasonable distribution. This is done by log compression, which is a histogram equalization process that rescales the signal histogram using a logarithmic function expressed as:

$$y = \frac{\log_{10}(1 + \alpha x)}{\log_{10}(1 + \alpha)}$$

where  $x$  is the input signal,  $\alpha$  is the compression factor that can be tuned, and  $y$  is the output. Here the input  $x$  is essentially the detected envelope from the previous step.

- Scan Conversion

The last step in the pipeline is to reshape the scan lines to increase resolution for display. Since all the previous processes focus mainly on obtaining reasonable resolution along the longitudinal dimension, scan conversion is the step focusing on increasing the lateral resolution. In this experiment, we upsample the lateral view by 4 to make lateral resolution comparable with the longitudinal, and the result will be the final image readied to display.

- Evaluation

To quantitatively examine the spatiotemporal resolution of our implemented reconstruction, we use peak-signal-to-noise-ratio (PSNR) and frame rate to evaluate the results. The frame rate is simply 1 over the total processing time which is evaluated from the beginning of beamforming to the end of scan convertor. PSNR is defined as:

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}$$

where MSE is the mean square error between the test image  $X$  and reference  $I_{ref}$ , with the same total number of pixel  $N$ :

$$MSE = \frac{1}{N} \sum_{i=1}^N (X(i) - I_{ref}(i))^2$$

All the implementation and evaluation are done in MATLAB with k-wave toolbox installed.

## 2.2 Convolutional Neural Network for Vagus Nerve Tracking

### 2.2.1 Faster R-CNN Model

Multiple Computer Vision algorithms implement different deep neural networks to accomplish object detection and classification. One of the most well-known networks for real-time detection is Faster Region-based Convolutional Neural Network (Faster R-CNN). Derived from R-CNN and Fast R-CNN,

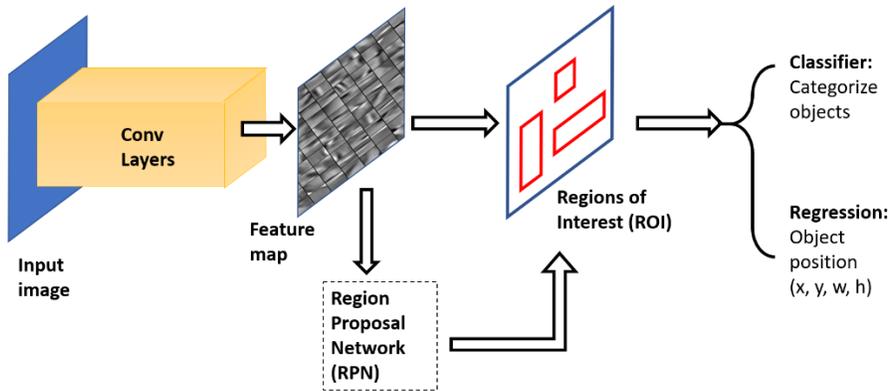


Figure 2.3 Faster R-CNN model with Region Proposal Network

this latest version beats them dramatically in processing speed for its shared convolutional layers from classifier network with the Region Proposal Network [30]. As demonstrated in the figure, the input image is fed into the shared convolutional layers to generate a feature map. The Region Proposal Network (RPN) then takes the feature map to compute for regions that have high probabilities to contain objects. The computed region proposals, together with the feature map, are sent to the region-of-interest pooling that finally generates the object class and the bounding box located at the object.

In this experiment, we use a template trainer in MATLAB to train our model. The trainer initializes the model with a regular CNN classifier. Here we try with two predefined models, “AlexNet”

and “VGG16” [31] [32]. The architecture for the two models in MATLAB is displayed in the figure 2.4, with a shallower and the other deeper network which will show up in the difference of the processing time [33]. We choose stochastic gradient descent with momentum for back propagation update method and use 0.001 for learning rate. The trainer starts with the classifier and adds the RPN accordingly while it trains the network.

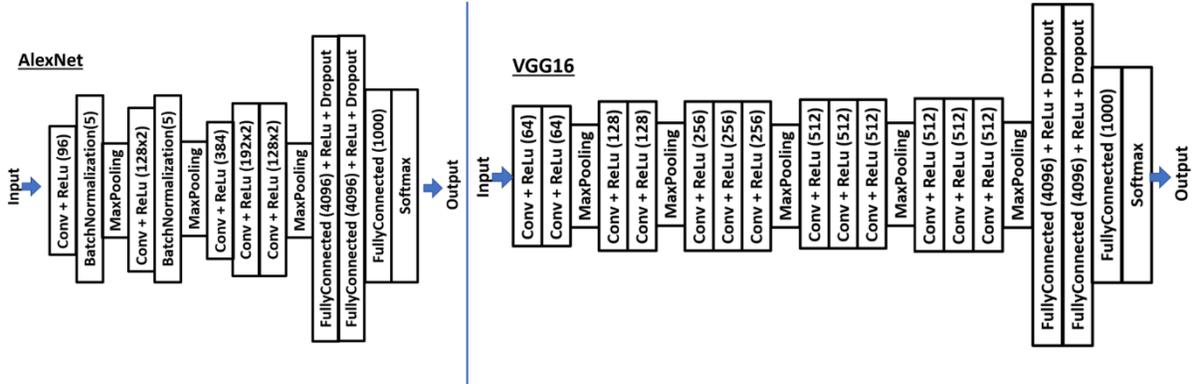


Figure 2.4 Architecture for AlexNet and VGGNet in MATLAB

## 2.2.2 Training Data Preparation

We use the L12-3v probe to get B-mode neck scans of our lab members and label the coordinate of the vagus nerve in each scan for ground truth with the ultrasound imaging expert assistance (Dr. Imanuel Lerman). We acquire 1091 images in total, including both left and right sides of the neck. We take the advantages of some prior knowledge about the vagus nerve for better data preparation. First, there are two sets of carotid artery + vagus nerve in our human neck, of which left and right side are approximately symmetrical. Therefore, it is reasonable to simply flip the images left-to-right for the sake of dataset augmentation. In this way, 2182 images become available for training and testing. Second, the vagus nerve is supposed to sit 1.2-2cm beneath the skin. The variations come from different thickness of a subject’s cervical tissue (skin, muscle and fat), yet it should never appear in close approximation to the skin surface or deeper inside the neck, i.e. below the carotid artery within the Longus Colli. This allows us to crop the reconstructed B-mode image to a smaller range of depth (1cm-2.5cm) so that the CNN

model can deal with the input image with much fewer pixels and thus dramatically decreasing processing time.

### 2.2.3 Proposed Image Guidance Pipeline

Current CAD techniques employ image detection algorithms on formed images that are already displayed on the screen. Therefore, current CAD techniques employ post-processing blocks and are thus not considered as real-time.

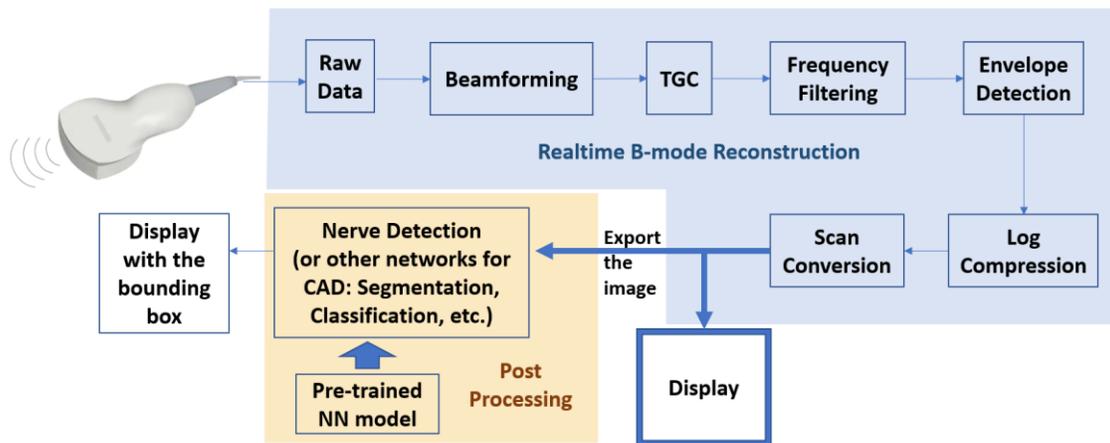


Figure 2.5 Conventional ultrasound image processing pipeline for CAD applications

In our experiment, we propose to apply the detection method before the final image so that the tracking information (i.e. bounding box around the nerve) can be shown together with the image on the screen and updated along the real-time scanning.

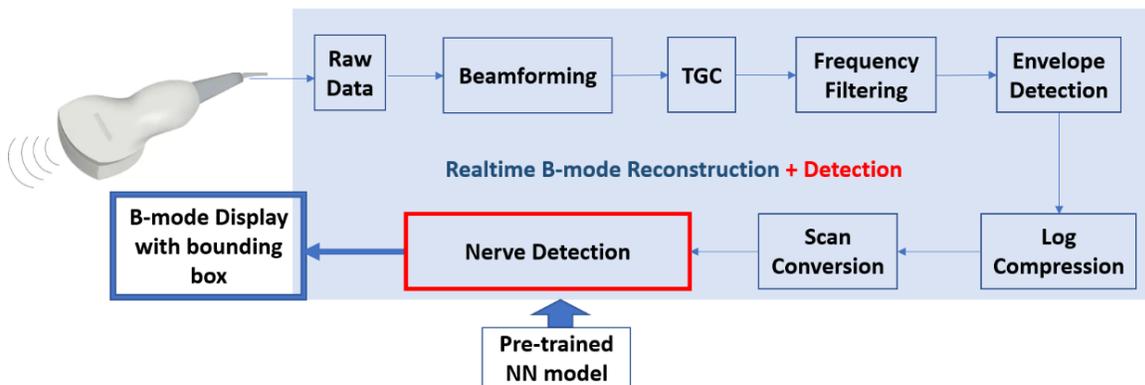


Figure 2.6 Proposed processing pipeline with detection network

## Chapter 3 Results

### 3.1 B-mode Image Reconstruction using Beamforming

#### 3.1.1 Beamforming Pipeline Results: Step-to-step

We apply our implemented reconstruction method to a real neck scan using the linear probe L12-3v and obtain the outputs for each step in the pipeline (figure 3.1). Figure 3.1a) shows the raw data over all 192 elements, and figure 3.1b) is the scan lines after the delay-and-sum beamforming. The significance of having beamforming as the very first step is illustrated clearly with these two results. In figure 3.1a) the tissue pattern is undistinguishable and sparse before the correction, whereas in figure 3.1b) we can roughly identify some smooth patterns. Therefore, beamforming provides the accurate structure and allows all the following steps to work on the correct basis. Figure 3.1c) is the result after Time Gain Compensation where more far-field patterns become visible. After this attenuation correction, the pixel intensity is boosted as we can tell from the dynamic range on the color bar. Figure 3.1d) shows the result after envelope detection, and figure 3.1e) is the result after logarithmic histogram equalization. Up to this point, we are able to get the overall structure for our field of interest. The last step is scan conversion which simply upsamples the data by 4 in lateral direction. Since the raw data gives the echo information up to its maximum acquisition depth (in this case is 52mm for L12-3v), we can select how deep we want for the final display, such as 38mm as shown in figure 3.1f). For comparison, the image generated from Verasonics built-in function is also displayed in figure 3.1g). We can see that our reconstructed result from the pipeline implementation is almost identical to the Verasonics computed image. In the following section, we carefully compare our reconstruction result to the Verasonics reference by computing the PSNR and cross-comparing among different numbers of elements for beamforming.

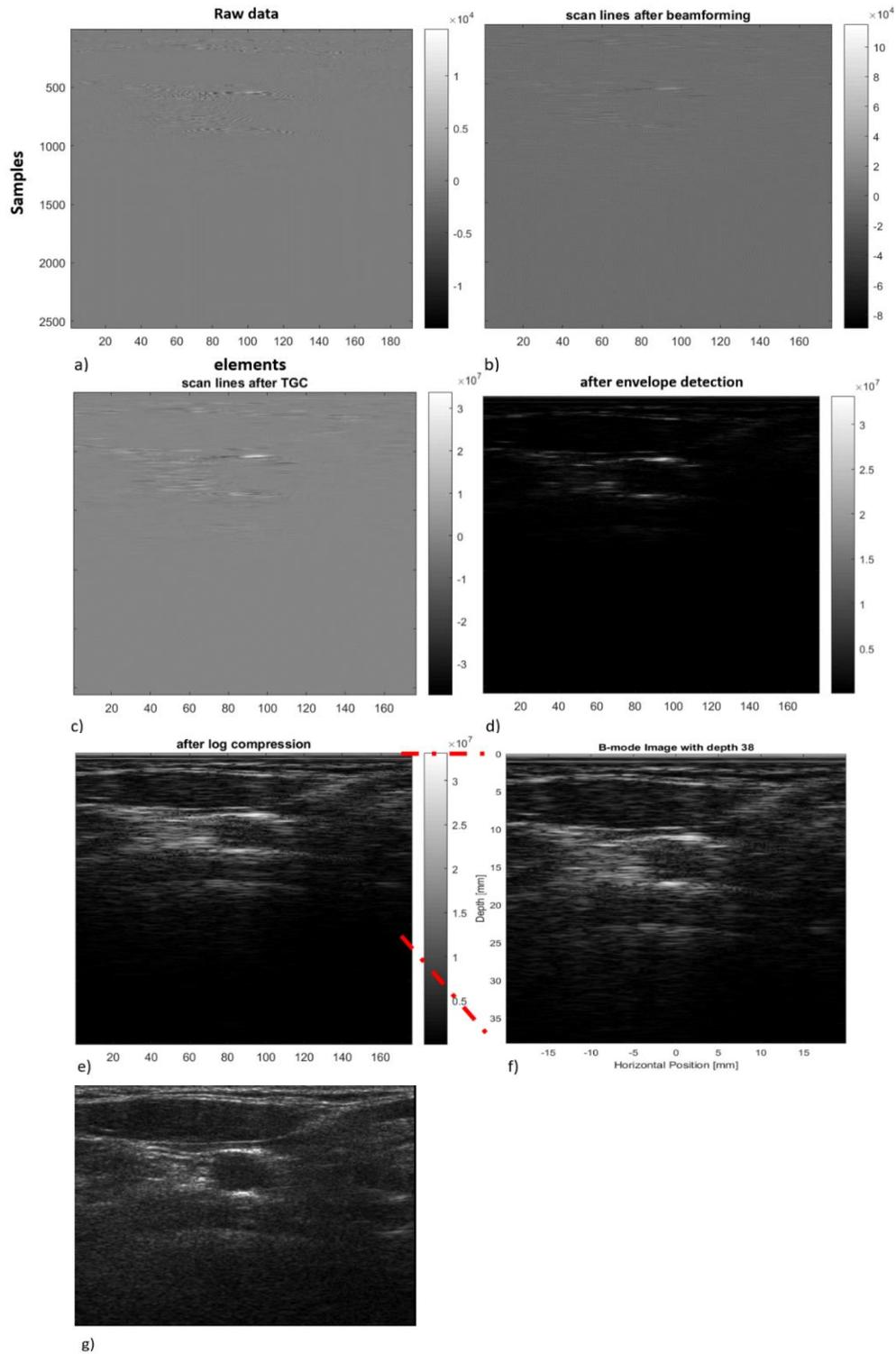


Figure 3.1 Step-by-step outputs from B-mode image reconstruction with beamforming: a) Raw data, and output from b) beamforming, c) TGC, d) envelope detection, e) log compression, f) scan conversion, and g) the reference image from Verasonics

### 3.1.2 Beamforming with Different Number of Elements

In this section, we examine carefully how numbers of elements for beamforming can affect the quality of reconstructed B-mode image. As we are not able to determine the number of elements employed in the built in Verasonics image reconstruction method we consider its result to be our reference for all the images we obtain from the pipeline. Verasonics provides a simulation mode, which shows the reconstruction of a virtual scan over some predefined patterns. For our experiment, we apply our implemented method both to the simulation mode and the actual scan on human neck. We examined the reconstruction time for one frame and the corresponding PSNR under different settings for beamforming (4/8/16/32 elements per sliding window). Under each setting, we take about 50~70 frames for 22mm, 30mm and 38mm of depth, respectively. The PSNR and frame rate are obtained by averaging among the frames for the same scenario.

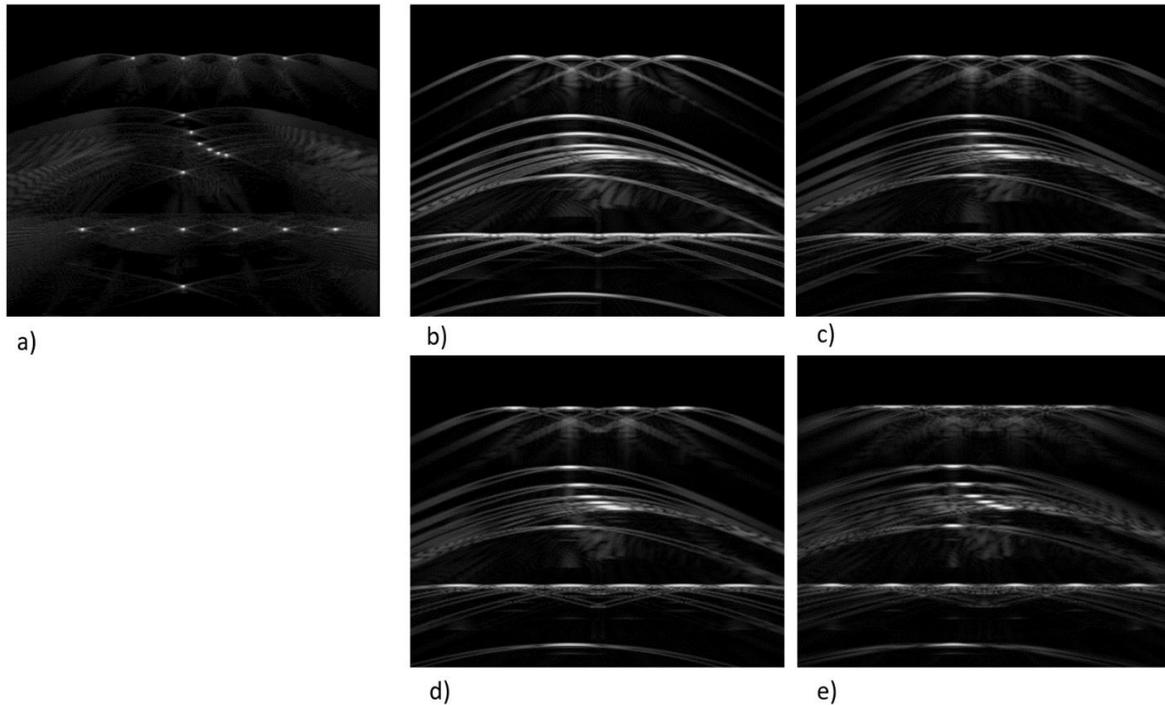


Figure 3.2 Reconstructed image under simulation mode using different number of elements for beamforming: a) reference from Verasonics, and the beamformed image with b) 4, c) 8, d) 16 and e) 32 elements per group.

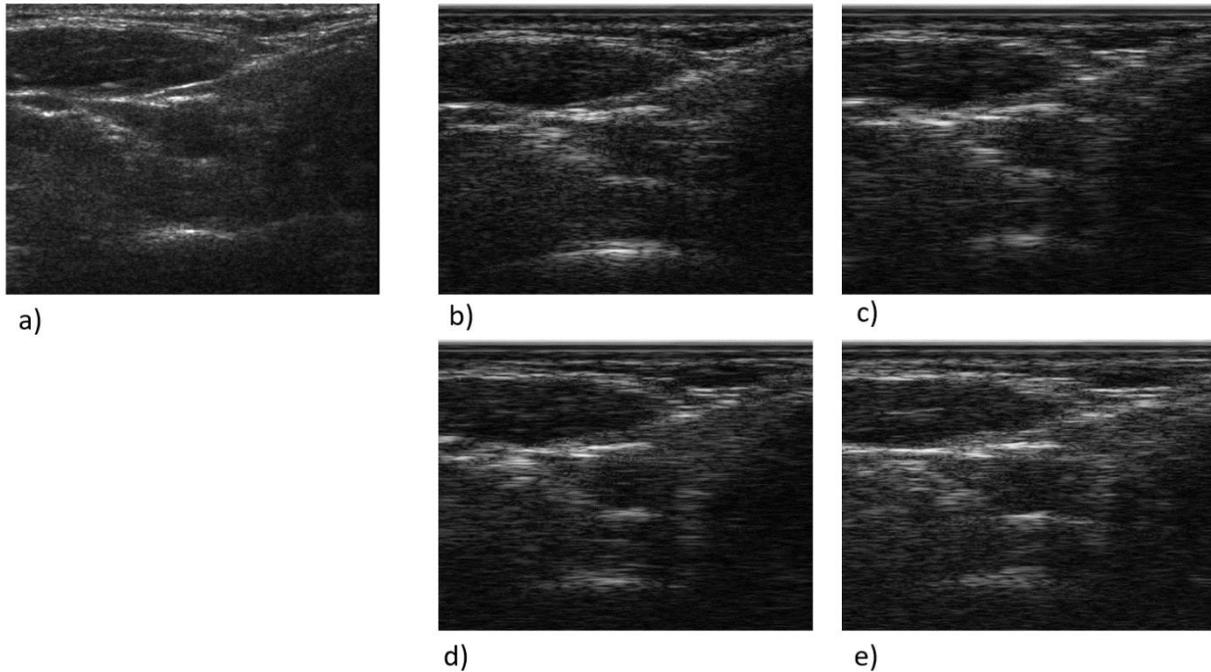


Figure 3.3 Reconstructed image under neck scan using different number of elements for beamforming: a) reference from Verasonics, and the beamformed image with b) 4, c) 8, d) 16 and e) 32 elements per group.

From the simulation mode in figure 3.2 we can clearly see that reconstruction becomes better (especially the far field patterns) as we include more elements per group in beamforming. It is slightly more difficult to differentiate the differences from the neck images in figure 3.3, since the textures get

Table 3.1: PSNR and frame rate comparison with different number of elements under simulation mode

number of elements	Evaluation	Depth (mm)			Average
		22	30	38	
4	Frame rate (Hz)	12.80	11.33	10.37	11.50
	PSNR (dB)	16.26	16.40	18.59	17.08
8	Frame rate (Hz)	12.33	11.04	10.05	11.14
	PSNR (dB)	17.58	17.81	20.29	18.56
16	Frame rate (Hz)	8.29	7.50	6.90	7.56
	PSNR (dB)	17.96	18.26	20.79	19.00
32	Frame rate (Hz)	5.26	4.90	4.48	4.88
	PSNR (dB)	18.53	18.90	21.38	19.60

Table 3.2: PSNR and frame rate comparison with different number of elements under neck scan

number of elements	Evaluation	Depth (mm)			Average
		22	30	38	
4	Frame rate (Hz)	12.64	11.27	10.14	11.35
	PSNR (dB)	14.56	16.33	18.98	16.62
8	Frame rate (Hz)	12.32	10.98	9.96	11.08
	PSNR (dB)	14.69	16.56	19.12	16.79
16	Frame rate (Hz)	8.31	7.50	6.89	7.56
	PSNR (dB)	14.45	16.55	18.91	16.64
32	Frame rate (Hz)	5.26	4.86	4.56	4.89
	PSNR (dB)	14.25	16.33	18.45	16.35

much more complex in the real scan than in simulation. However, we can see the trend from the two tables.

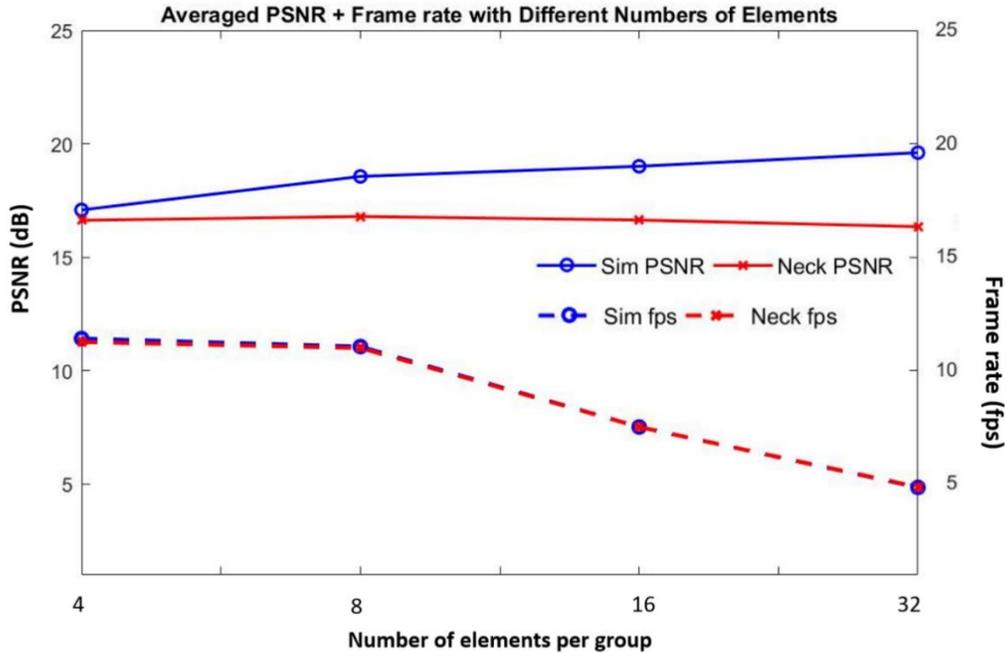


Figure 3.4 Averaged PSNR and frame rate with different number of elements. Frame rate depends on software implementation only. Since both simulation and real scan undergo the same computation, time consumption is identical. PSNR depends on the complexity of the data including the texture, sampling noise, echoic speckles, etc. Therefore, PSNR evaluation diverges between the two scenarios.

For both simulation mode and neck scan with same number of elements, image quality improves as we increase the acquisition depth, which makes perfect sense since more sample points are required. When we include more elements per group for beamforming, the computational cost gets increased as well, which explains the decreasing frame rate. The PSNR of the neck scan is not as significantly influenced as that of the simulation mode, yet there is still a measurable increase when using 8 elements instead of 4. The nominal increase in PSNR for 16 and 32 elements is possibly caused by fewer scan lines available from the beamforming, since the total number of elements is fixed, and we are grouping more channels for one scan line. Nevertheless, greater element number only result in small gains in PSNR fluctuations. In the summary plot shown as figure 3.4, we can clearly see the trade-off between number of elements and the reconstruction efficiency (image quality + frame rate). We aim for a relatively high frame rate with fair image quality adequate for use in our CNN. Considering both the simulation and real scan, we choose 8 and 16 elements for our main data collection that will eventually be used to train the neural network. 8-elements was notable for its low frame rate and minimal decrement in PSNR, and while the cost on the PSNR; with 16-elements gives a slightly higher PSNR there is a significantly slower frame rate. In sum, in our efforts for relative real time nerve identification and classification (that take into account movement due pulsation); we decreased our frame rate to allow more processing time allotted to the neural network for nerve detection.

### 3.2 Vagus Nerve Detection from Beamformed Data Using Faster R-CNN

Our faster R-CNN model was trained using the VGG16 and AlexNet on a Windows server with GPU capability (CUDA Device: Tesla K80, 12GB). 1309 images were used for training, and 873 images were tested. Given the ground truth provided by the expert medical doctor (Dr. Imanuel Lerman), the vagus nerve fibers are hypoechoic surrounded by hyperechoic fascicles and further by tissue within the carotid sheath that is easily identified with B-mode ultrasound imaging. In most cases the vagus nerve is

located between the carotid and jugular vein which is more ventrolateral and proximal toward the skin surface.

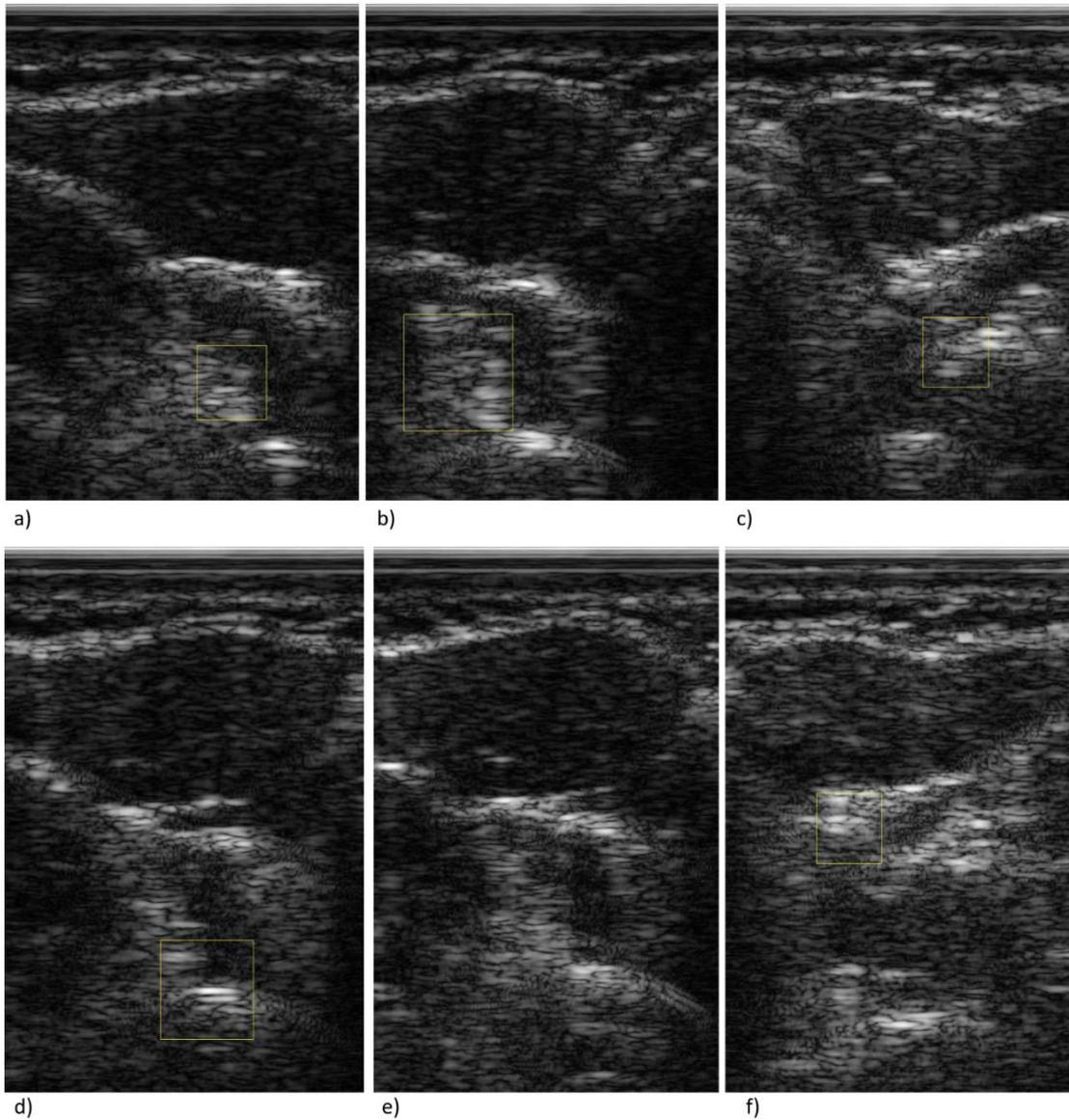


Figure 3.5 Vagus nerve detection results from beamformed data: correct results on the a) b) left side and c) right side of the neck; wrong results on the d) e) left side and f) right side of the neck.

Figure 3.5 a), b), c) show results obtained from the reconstructed B-mode neck scan (left, left and right side, respectively). Figure #d), e) and f) are some examples of failed detection, which yield a wrong location or fail to find the target at all within the scene. Note that the trained and tested images are cropped to 725x725 within the depth of 1cm~2.5cm only. The displayed images in figure # are the view from skin (0cm) to 2.5cm, which is the half view of the maximum depth.

The performance of the trained vagus nerve detectors are shown in the following table (table 3.3). AlexNet has fewer layers, so it requires less time than VGG16 to finish training. It also has a faster frame rate of 6.16Hz (162ms per frame), compared to 2.41Hz (415ms) from VGG16. A shorter frame processing time is critically needed as we aim to achieve real-time location updates with the detector

Table 3.3: Performance of vagus nerve detector with two models

	AlexNet	VGG16
<b>Training Time (1309 images, 725x725)</b>	<b>54mins, 49s</b>	<b>4hrs, 9mins, 29s</b>
<b>Testing (873 images, 725x725)</b>		
<b>Average precision</b>	<b>0.46</b>	<b>0.54</b>
<b>Average processing time</b>	<b>162.4ms</b>	<b>414.9ms</b>

combined with the B-mode reconstruction. However, VGG16 yields a slightly higher accuracy (54%) than AlexNet (46%) with a deeper network. The average precision from the tested images is roughly 50%, implying that in general this is a weak detector.

The low accuracy can be caused by the resolution of the B-mode reconstructed images. Therefore, we induce the training process with 796 additional images acquired from Philips Healthcare® medical imaging ultrasound device CX50 utilizing the 192 element 12-3MHz linear array transducer.

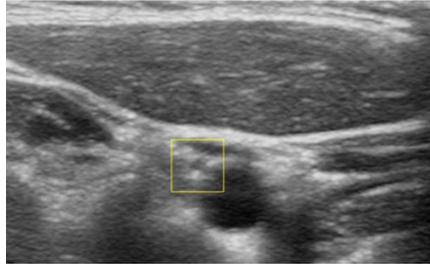


Figure 3.6: B-mode scan using clinical device from Philips Healthcare®.

They provide devices for clinical diagnosis that has much higher resolution. Although the new 796 images are reconstructed externally, not from the beamformer in this work, they serve only as a substrate for the trainer with the aim to improve the performance. We then applied the CX50 derived updated detector among the same testing images from our Verasonics 12-3MHz beamforming reconstruction, with results are shown in table 3.4. With better training data, the frame rate did not significantly change.

Table 3.4: Performance of vagus nerve detectors (including external training data)

	AlexNet		VGG16	
	1309 beamformed images	+796 Philips data	1309 beamformed images	+796 Philips data
Training Time	54mins, 49s	1hr, 8mins, 53s	4hrs, 9mins, 29s	4hrs, 35mins, 20s
<b>Testing (873 beamformed images)</b>				
Average precision	0.46	<b>0.35</b>	0.54	<b>0.63</b>
Average processing time	162.4ms	<b>180.5ms</b>	414.9ms	<b>418.9ms</b>

Examining accuracy, VGG16 improved by 9%, which is what we expected, but AlexNet showed a decrement in performance. This is understandable considering the development of those nets. AlexNet is ecologically an older CNN when compared to VGG16, that contains inherent problems impeding its capability to resolve deeper data sets as it was developed for shallow networks. VGG16, as a deeper network, updated and fixed some issues for better application. Therefore, at this time VGG16 is a showed some promise as an initial model.

It requires expert medical knowledge to identify the vagus nerve from the surrounding tissue in a B-mode scan. The size of the nerve is very small compared to carotid artery, while the vagus nerve is likely to have spatial similarity at some other interface of tissues within the image that may mislead the detection. In general, it is a much harder problem to detect the vagus nerve in B-mode scan than to detect a normal object in an RGB image. Of particular note the current image detection algorithm relies heavily on the reconstruction of the beamformed data. Therefore, improvement of the beamforming implementation is the logical next step for our future.

## Chapter 4 Summary and Future Work

### Conclusion:

In this work, we achieve the implementation of beamforming and scan conversion for a research-grade B-mode ultrasound imaging system. We implement a major innovation; in that we are able to detect and track the vagus nerve, in real-time, by inserting the Faster R-CNN based nerve detection model into the beamforming pipeline. To achieve more efficient real-time tracking of the nerve from B-mode images, we need to speed up both the B-mode reconstruction and detection network. Future work will build on prior beamforming implementation as well as focus on improved design and validation of the neural network.

### Future Work:

In beamforming, time shift and summation using loops can be replaced by a simple matrix manipulation using linear algebra to reduce computational complexity. Variable number of elements controlled by directivity should also be considered, since beam directivity provides additional weights that determine how much the echoes contribute at each angle of the elements (i.e. the one that's perpendicular weighs more than the one at the far side). Moreover, the current work only applies to use of a linear probe, while future work will focus on curved array, phased array and other configurations.

As is evident from our current detection rate of 50-63% we will need to re-design the nerve detection network for higher accuracy. Instead of having the filters to extract the common features such as edges, the layers should be re-organized and modified to approach the identification of the vagus nerve in a B-mode image. First, we plan to identify where the carotid artery is, and then search edges of the carotid with built in search algorithm aimed to ventrolateral surface of the carotid (i.e., known location of vagus nerve) and not the medial or dorsal in a deeper view. With a more reasonable architecture to detect the nerve, we anticipate that we will improve the CNN accuracy as well as computational efficiency. Moreover, the current well-organized layers are functional, with some redundant layers noted that can

subsequently be discarded to avoid detouring computations and thus saves processing time. As is well established: 1) expansion of training dataset including more scans from multiple other subjects and 2) injecting redundant artificial noise will make the detector more robust.

The main thrust of this thesis was to obtain real time vagus nerve detection as it moves in sync with the carotid pulsation. Further future work will also focus on location updates for movement compensation due to patient and or equipment movement. As mentioned in the introduction, from the detector we have the nerve coordinates to guide the focused ultrasound stimulus, but there are unavoidable movements that can shift the target away from the stimulus beam which is unsafe. Therefore, we need to look at the difference of the nerve coordinates between frames and consider macro and micro movements as we update the location to the system. A micro movement comes from the periodic pulsation of carotid artery, and it can be compensated by the beam steering capability of the stimulating transducer. The custom designed multi-element random array from Sonic Concept Inc. can steer the beam

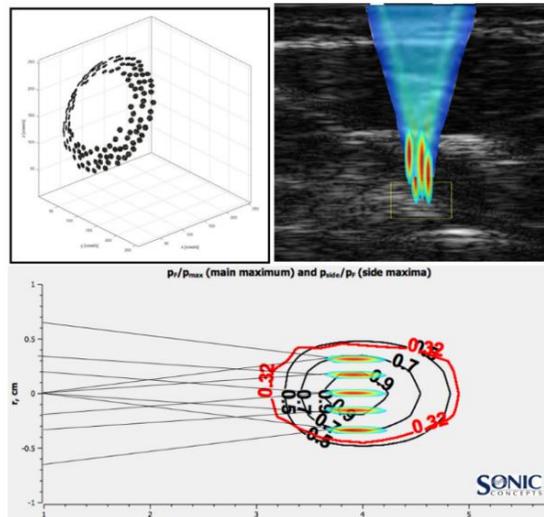


Figure 4.1 Beam steering capability to compensate focus to a certain distance away from the original center. A macro movement, instead, is possibly caused by the patient or displacement of the imaging probe. This information will be sent to robotic arm that holds the transducer for macro-movement position correction. The UR3 programmable robotic arm from Universal Robotics can take serial control through python codes and adjust the end position in x, y and z directions. We've already proved this idea to be working by sending a sequence of moving coordinates in

a demo video to control the position of the arm. We are planning to purchase a higher version, UR3e, that has frame rate of 500Hz (UR3 125Hz) so that it is sufficient for movement updates on millisecond scale. Therefore, the future work will include having this system attach to the end of the B-mode reconstruction and detection pipeline so that the nerve location will be used to determine movements so as to control for either beam steering or robotic arm positioning.

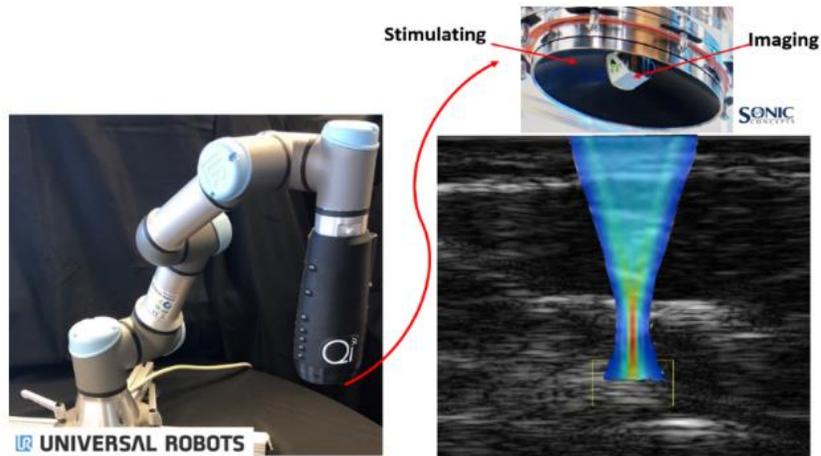


Figure 4.2 Robotic arm programming to compensate large-scale movements

## References

- [1] I. Lerman, R. Hauger, L. Sorkin, J. Proudfoot, B. Davis, A. Huang, K. Lam, B. Simon and D. G. Baker, "Noninvasive transcutaneous vagus nerve stimulation decreases whole blood culture-derived cytokines and chemokines: a randomized, blinded, healthy control pilot trial," *Neuromodulation: Technology at the Neural Interface*, pp. 283-290, 2016.
- [2] A. P. Mourdoukoutas, D. Q. Truong, D. K. Adair, B. J. Simon and M. Bilkson, "High-Resolution Multi-Scale Computational Model for Non-Invasive Cervical Vagus Nerve Stimulation," *Neuromodulation: Technology at the Neural Interface*, pp. 261-268, 2018.
- [3] M. Ter Laan, J. M. C. Van Dijk, J. W. J. Elting, M. J. Staal and A. R. Absalom, "Sympathetic regulation of cerebral blood flow in humans: a review," *British journal of anaesthesia*, vol. 111, no. 3, pp. 361-367, 2013.
- [4] M. Swissa, S. Zhou, I. Gonzalez-Gomez, C.-M. Chang, A. C. Lai, A. W. Cates, M. C. Fishbein, H. S. Karagueuzian, P.-S. Chen and L. S. Chen, "Long-term subthreshold electrical stimulation of the left stellate ganglion and a canine model of sudden cardiac death," *Journal of the American College of Cardiology*, vol. 43, no. 5, pp. 858-864, 2004.
- [5] W. J. Tyler, S. W. Lani and G. M. Hwang, "Ultrasonic modulation of neural circuit activity," *Current opinion in neurobiology*, pp. 222-231, 2018.
- [6] L. R. Gavrilov, G. V. Gersuni, O. B. Ilyinsky, M. G. Sirotyuk, E. M. Tsirulnikov and E. E. Shchekanov, "The effect of focused ultrasound on the skin and deep nerve structures of man and animal," *In Progress in brain research*, vol. 43, no. Elsevier, pp. 279-292, 1976.
- [7] W. Legon, T. F. Sato, A. Opitz, J. Mueller, A. Barbour, A. Williams and W. J. Tyler, "Transcranial focused ultrasound modulates the activity of primary somatosensory cortex in humans," *Nature neuroscience*, p. 322, 2014.
- [8] L. J. Noble, I. J. Gonzalez, V. B. Meruva, K. A. Callahan, B. D. Belfort, K. R. Ramanathan, E. Meyers, M. P. Kilgard, R. L. Rennaker and C. K. McIntyre, "Effects of vagus nerve stimulation on extinction of conditioned fear and post-traumatic stress disorder symptoms in rats," *Translational psychiatry*, vol. 7, no. 8, p. e1217, 2017.
- [9] F. A. Koopman, S. S. Chavan, S. Miljko, S. Grazio, S. Sokolovic, P. R. Schuurman, A. D. Mehta, Y. A. Levine, M. Faltys, R. Zitnik and K. J. Tracey, "Vagus nerve stimulation inhibits cytokine production and attenuates disease severity in rheumatoid arthritis," *National Academy of Sciences*, vol. 113, no. 29, pp. 8284-8289, 2016.
- [10] J. W. Salatino, K. A. Ludwig, T. D. Kozai and E. K. Purcell, "Glial responses to implanted electrodes in the brain," *Nature biomedical engineering*, vol. 1, no. 11, p. 862, 2017.
- [11] W. R. Hedrick, D. L. Hykes and D. E. Starchman, *Ultrasound physics and instrumentation*, 2005.

- [12] T. Wagner, A. Valero-Cabre and A. Pascual-Leone, "Noninvasive human brain stimulation.," *Annu. Rev. Biomed. Eng.*, vol. 9, pp. 527-565, 2007.
- [13] W. J. Tyler, Y. Tufail, M. Finsterwald, M. L. Tauchmann, E. J. Olson and C. Majestic, "Remote excitation of neuronal circuits using low-intensity, low-frequency ultrasound," *PLoS one*, vol. 3, no. 10, p. e3511, 2008.
- [14] Y. Tufail, A. Matyushov, N. Baldwin, M. L. Tauchmann, J. Georges, A. Yoshihiro and S. I. H. Tillery, "Transcranial pulsed ultrasound stimulates intact brain circuits," *Neuron*, vol. 66, no. 5, pp. 681-694, 2010.
- [15] R. L. King, J. R. Brown, W. T. Newsome and K. B. Pauly, "Effective parameters for ultrasound-induced in vivo neurostimulation," *Ultrasound in medicine & biology*, vol. 39, no. 2, pp. 312-331, 2013.
- [16] A. Liu, L. Song, L. Li, X. Wang, H. Lin and Y. Wang, "A controlled trial of transcutaneous vagus nerve stimulation for the treatment of pharmacoresistant epilepsy," *Epilepsy & Behavior*, vol. 39, pp. 105-110, 2014.
- [17] J. Redgrave, D. Day, H. Leung, P. J. Laud, A. Ali, R. Lindert and A. Majid, "Safety and tolerability of Transcutaneous Vagus Nerve stimulation in humans; a systematic review," *Brain stimulation*, vol. 11, no. 6, pp. 1225-1238, 2018.
- [18] E. J. Juan, R. González, G. Albors, M. P. Ward and P. Irazoqui, "Vagus nerve modulation using focused pulsed ultrasound: potential applications and preliminary observations in a rat," *International journal of imaging systems and technology*, vol. 24, no. 1, pp. 67-71, 2014.
- [19] K. M. Wasilczuk, K. C. Bayer, J. P. Somann, G. O. Albors, J. Sturgis, L. T. Lyle, J. P. Robinson and P. P. Irazoqui, "Modulating the Inflammatory Reflex in Rats Using Low-Intensity Focused Ultrasound Stimulation of the Vagus Nerve," *Ultrasound in medicine & biology*, vol. 45, no. 2, pp. 481-489, 2019.
- [20] T. Y. Wu, N. Guo, C. Y. Teh and J. X. Hay, "Theory and fundamentals of ultrasound," in *Advances in ultrasound technology for environmental remediation*, 2013, pp. 5-12.
- [21] O. Hadjerci, A. Hafiane, P. Makris, D. Conte, P. Vieyres and A. Delbos, "Nerve detection in ultrasound images using median gabor binary pattern," *International Conference Image Analysis and Recognition*, pp. 132-140, 2014.
- [22] C. N. Ta, "Computer-Aided Diagnosis of Tumors with Contrast-Enhanced Ultrasound," *Doctoral dissertation, UC San Diego*, 2015.
- [23] O. Hadjerci, A. Hafiane, D. Conte, P. Makris, P. Vieyres and A. Delbos, "Computer-aided detection system for nerve identification using ultrasound images: A comparative study," *Informatics in Medicine Unlocked*, vol. 3, pp. 29-43, 2016.
- [24] Y. LeCun, K. Kavukcuoglu and C. Farabet, "Convolutional networks and applications in vision," *IEEE International Symposium on Circuits and Systems*, pp. 253-256, 2010.

- [25] Y. Bar, I. Diamant, L. Wolf and H. Greenspan, "Deep learning with non-medical training used for chest pathology identification," *Medical Imaging 2015: Computer-Aided Diagnosis*, vol. 9414, no. International Society for Optics and Photonics, p. 94140V, 2015.
- [26] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1207-1216, 2016.
- [27] E. Smistad and L. Løvstakken, "Vessel detection in ultrasound images using deep convolutional neural networks," *Deep Learning and Data Labeling for Medical Applications*, pp. 30-38, 2016.
- [28] K. Lekadir, A. Galimzianova, À. Betriu, M. del Mar Vila, L. Igual, D. L. Rubin, E. Fernández, P. Radeva and S. Napel, "A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound," *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, pp. 48-55, 2016.
- [29] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, p. 021314, 2010.
- [30] S. Ren, K. He, R. Girshick and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, pp. 91-99, 2015.
- [31] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097-1105, 2012.
- [32] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*.
- [33] H. Kataoka, K. Iwata and Y. Satoh, "Feature evaluation of deep convolutional neural networks for object recognition and detection," *arXiv preprint arXiv:1509.07627*, 2015.